Available online at www.sciencedirect.com
SCIENCE DIRECT®
Plasmid 52 (2004) 13–30
PLASMID
www.elsevier.com/locate/yplas

Replication of a unit-copy plasmid F in the bacterial cell cycle: a replication rate function analysis

Paul F. Morrison[a] and Dhruba K. Chattoraj[b,*]

[a] *Division of Bioengineering and Physical Science, ORS, National Institutes of Health, Bethesda, MD 20892-5766, USA*
[b] *Laboratory of Biochemistry, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892-4255, USA*

Received 8 January 2004, revised 4 April 2004
Available online 6 May 2004

Abstract

For stability, the replication of unit-copy plasmids ought to occur by a highly controlled process. We have characterized the replication dynamics of a unit-copy plasmid F by a replication rate function defined as the probability per unit age interval of the cell cycle that a plasmid will initiate replication. Analysis of baby-machine data [J. Bacteriol. 170 (1988) 1380; J. Bacteriol. 179 (1997) 1393] by stochastics that make no detailed reference to underlying mechanism revealed that this rate function increased monotonically over the cell cycle with rapid increase near cell division. This feature is highly suggestive of a replication control mechanism that is designed to force most plasmids to replicate before cells undergo division. The replication rate function is developed anew from a mechanistic model incorporating the hypotheses that initiators are limiting and that steric hindrance of origins by handcuffing control initiation of replication. The model is based on correctly folded initiator protein monomers arising from an inactive dimer pool via chaperones in limiting amounts, their random distribution to high affinity sites (iterons) at the origin (*ori*) and an outside locus (*incC*), the statistical mechanics of bound monomer participation in pairing the two loci (*cis*-handcuffing), and initiation probability as proportional to the number of non-handcuffed *ori*-saturated plasmids. Provided *cis*-handcuffing is present, this model closely accounts for the shape of the replication rate function derived from experiment, and reproduces the observation that replication occurs throughout the cell cycle. Present concepts of iteron-based molecular mechanisms thus appear capable of yielding a quantitative description of unit-copy-number plasmid replication dynamics.
Published by Elsevier Inc.

1. Introduction

The replication of unit-copy plasmids is a unique event in the cell cycle. In slow growing *Escherichia coli* (with a generation time of an hour or more), cells are born with just one copy of the plasmid that replicates only once with high probability before the cell divides (Austin and Eichorn, 1992). The control of replication must therefore be stringent as well as efficient. Evidence of such stringent control has been reported for unit-copy plasmid F, where experimental high copy number deviations appear to invoke such strong negative feedback that replication appears to be totally

* Corresponding author. Fax: 1-301-480-1493.
E-mail address: chattord@dc37a.nci.nih.gov (D.K. Chattoraj).

0147-619X/$ - see front matter. Published by Elsevier Inc.
doi:10.1016/j.plasmid.2004.04.001

switched off at only twice the steady state copy number (Tsutsui and Matsubara, 1981). The situation seems to be similar for another unit-copy plasmid P1, where the plasmid fails to transform a cell that already has one chromosomally integrated plasmid copy (Pal and Chattoraj, 1988).

The molecular mechanism by which such control is exerted has been the subject of intense study. Broad outlines of the processes involved have emerged for several plasmids, most notably for F, P1, and RK2 (Chattoraj, 2000). These plasmids all code for initiator proteins that auto-repress or otherwise control their own production, as well as bind to two sets of repeats termed iterons, one set of which in the origin (*ori*) when saturated becomes responsible for initiating plasmid replication. The other set (called *incC* in F) acts solely as a negative regulator of initiation frequency. In the specific case of the F plasmid, the initiator protein RepE suppresses its own production when dimeric but binds to iterons when monomeric. Direct observation by X-ray diffraction has confirmed the binding of these monomers to their iteron targets (Komori et al., 1999). Tsutsui et al. (1983) hypothesized that RepE is limiting for initiation and its titration by *incC* iterons delays saturation of *ori* iterons and thereby, the timing of plasmid initiation. Later investigators observed that this cannot be the sole process involved since replication frequency did not increase in the presence of excess initiator protein in all the three plasmids P1, RK2, and F (Durland and Helinski, 1990; Pal and Chattoraj, 1988; Uga et al., 1999). Electron microscopy and other biochemical analyses revealed that the initiator proteins could interact when bound to two different iteron sets, either in *cis* or *trans*, in a process termed handcuffing (Kittell and Helinski, 1991; Mukherjee et al., 1985; Pal and Chattoraj, 1988; Uga et al., 1999). These interactions appeared responsible for suppressing additional rounds of initiation in the same cell cycle should the initiators become excessive.

The identification of the molecular components involved in unit-copy-number plasmid replication has now advanced to the point, where one can begin to quantitatively characterize the dynamics of their interaction, along the way testing prevailing concepts for self-consistency and completeness. In this article, we focus on the results of the 'baby-machine' experiments that probe these dynamics in the cell cycle and attempt to find kinetic explanations for the observed behavior (Helmstetter et al., 1997; Leonard and Helmstetter, 1988). We first show that a mathematical expression for the probability of a plasmid's replication per unit time as a function of age in the cell cycle, $\lambda(a)$ (henceforth termed the replication rate function), can be generated from the baby-machine experimental data, without reference to any specific underlying molecular mechanism. We then develop this function anew from a mechanistic model incorporating present hypotheses of initiator-titration and iteron-handcuffing that control initiation of plasmid replication. Finally we perform a test of this mechanism by assessing whether, and under what conditions, the titration-handcuffing model is capable of yielding a $\lambda(a)$ function that is in quantitative agreement with that obtained directly from experiment.

2. Analysis

2.1. Stochastic formulation of unit-copy-number plasmid replication

The simplest phenomenological approach to a description of unit-copy-number plasmid replication begins with a stochastic description of the process in a cell population. In this section, we formulate the replication in terms of an age-dependent replication rate function $\lambda(a)$ and then show that an analytical expression for it can be derived from baby-machine experimental data with some simplifying assumptions.

First, consider an asynchronous host population of bacterial cells in exponential growth containing 1–2 copies of a plasmid. Within this population, cohorts of cells exist at various ages with those of youngest age (just post-cell division) containing a single plasmid and those of oldest age (just prior to the next cell division) containing two plasmids. At intermediate ages, an age cohort consists of a mix of one- and two-copy-containing cells depending on whether or not plasmid repli-

cation has taken place in a given cell. The probability of finding a unit-copy plasmid within a particular age cohort is defined as $p_1(a)$ where a is the age measured in time from the last cell division. This probability density evolves with age according to the stochastic equation

$$\frac{\mathrm{d}p_1(a)}{\mathrm{d}a} = -\lambda(a)p_1(a), \quad p_1(0) = 1, \tag{1}$$

i.e., given the probability $p_1(a)$ of selecting a unit-copy cell from an age cohort, there is a conditional probability per unit time ($\lambda(a)$) that its plasmid will then replicate. $\lambda(a)$ is defined as the replication rate function. Eq. (1) may be formally integrated for $p_1(a)$ yielding

$$p_1(a) = \exp\left[-\int_0^a \lambda(a')\,\mathrm{d}a'\right]. \tag{2}$$

Depending upon the exact form of $\lambda(a)$, plasmid replication may appear to be more or less cycle specific (unpublished results).

It is possible to extract the exact form of $\lambda(a)$, and implicitly, information about the replication mechanism on which it depends, from the data generated by baby-machine experiments. Briefly, these experiments (Helmstetter et al., 1997; Leonard and Helmstetter, 1988) involve pulse labeling an asynchronous population of plasmid-containing host cells with [^{3}H]thymidine, attachment of these cells to a nitrocellulose filter under growth permissive conditions, and age cohort collection of newborn cells that are continuously released from the filter. Following autoradiographic or scintillation analysis, the average plasmid-radioactivity per cell collected is reported for each cohort. The first daughter cells eluted from the filter are progeny of the cells attached to the filter at the end of the division cycle, and the last to be eluted in a generation cycle are those that are the progeny of cells attached to the filter near the beginning of the cycle.

Hence, denoting the bacterial cell cycle time as T and the age of the cell as a, the time of release of daughter cells from the filter, t_{ex}, is related to their age at labeling by

$$t_{\mathrm{ex}} = T - a. \tag{3}$$

Next, we observe that the rate at which *labeled* 2-copy cells are generated at a given age during short-term pulse labeling is

$$\frac{\partial n_2^*(t,a)}{\partial t} = \lambda(a)n_1(t,a), \tag{4}$$

where at age a and chronological time t, $n_2^*(t,a)$ is the number density of *labeled* 2-copy cells per age interval and $n_1(t,a)$ is the number density of *unlabeled* unit-copy cells. (If the pulse period is not approximately instantaneous, then Eq. (4) must be replaced by a continuity equation that also accounts for changes in the number density function resulting from cell aging during the pulse duration.) An expression for $n_2^*(t,a)$ can be obtained by integrating Eq. (4) over the pulse duration Δ and noting that the number density n_2^* is initially zero. Hence

$$n_2^*(t,a) \approx \lambda(a)n_1(t,a)\Delta \tag{5}$$

$$= \lambda(a)p_1(a)n(t,a)\Delta. \tag{6}$$

In the last line, the definition of $p_1(a)$ as the fraction of unit-copy cells in an age cohort has been employed to re-express the unit-copy density $n_1(t,a)$ in terms of the total cell density $n(t,a)$.

In the cell collection phase of the baby-machine experiments, cohorts of cells that progress through cell division are gathered over a short time interval Δ_e. The cells being collected, thus correspond to the flux of cells moving across the $a = T$ boundary for time Δ_e at a maturation velocity $\mathrm{d}a/\mathrm{d}t$ of unity. Mathematically, the fluxes of labeled and total cells integrated over the collection time are $n_2^*(t,a)\Delta_e$ and $n(t,a)\Delta_e$, and the amount of radioactivity in the labeled group is $\gamma n_2^*(t,a)\Delta_e$, where γ is the fixed amount of radiolabel incorporated into a single plasmid's DNA during replication. The experimental variable reported as the average plasmid-radioactivity per cell collected, defined here as D, is thus

$$D = \frac{\gamma n_2^*(t,a)\Delta_e}{n(t,a)\Delta_e} = \frac{\gamma n_2^*(t,a)}{n(t,a)} = \gamma\lambda(a)p_1(a)\Delta, \tag{7}$$

where Eq. (6) has been substituted for $n_2^*(t,a)$ to yield the right hand side of Eq. (7).

To within a constant, D may be expressed solely in terms of the replication rate function, $\lambda(a)$. This

is achieved by taking the natural logarithm of Eq. (7) and substituting Eq. (2) for $p_1(a)$ to yield

$$\ln[D] = \ln[\gamma A] + \ln[\lambda(a)] - \int_0^a \lambda(a')\,\mathrm{d}a'. \tag{8}$$

Experimentally, $\ln[D]$ of miniF plasmid closely approximates linear behavior throughout the cell cycle at both short and long cell division times (Helmstetter et al., 1997; Leonard and Helmstetter, 1988) so that

$$\ln[D] = -At_{\mathrm{ex}} + B = -A(T-a) + B, \tag{9}$$

where Eq. (3) has been used for t_{ex}. This linearity also extends across several cell division times (reflecting the dilution of labeled plasmid DNA per collected cell due to continued cell division of the mother cells attached to the nitrocellulose membrane), so that

$$A = (\ln 2)/T. \tag{10}$$

An equation for the determination of the $\lambda(a)$ function follows from differentiating Eq. (8) for $\mathrm{d}\ln[D]/\mathrm{d}a$ and noting from Eq. (9) that this derivative is just A, i.e.,

$$\frac{\mathrm{d}\ln[D]}{\mathrm{d}a} = \frac{1}{\lambda(a)}\frac{\mathrm{d}\lambda(a)}{\mathrm{d}a} - \lambda(a) = A. \tag{11}$$

This differential equation is a Bernoulli equation that has the monotonically increasing solution (Matthews and Walker, 1965)

$$\lambda(a) = \frac{1}{\left(\frac{1}{\lambda_0} + \frac{1}{A}\right)\exp[-Aa] - \frac{1}{A}}, \tag{12}$$

where λ_0 is the value of λ at birth. λ_0 can be obtained from the observed probability ω of detecting a cell still containing only an unreplicated plasmid at cell division, i.e., from the condition

$$p_1(T) = \omega \tag{13}$$

combined with Eqs. (2), (10), and (12). ω may be considered as the inefficiency of unit-copy replication.In the *special case* of perfect 1- to 2-copy number control, ω is zero and $p_1(T) = 0$, i.e., no cells exist in which plasmid replication has not taken place. From Eq. (2), this is equivalent to the condition that

$$\int_0^T \lambda(a')\mathrm{d}a' = \infty. \tag{14}$$

This integral reaches infinity only when the denominator of Eq. (12) is zero, i.e., when

$$\lambda_0 = \frac{A\exp[-AT]}{1-\exp[-AT]}. \tag{15}$$

Substituting Eq. (15) for λ_0 and Eq. (10) for A into Eq. (12), the fractional replication rate function under perfect copy control is found to be

$$\lambda(a) = T^{-1}\ln 2(2^{1-a/T} - 1)^{-1}. \tag{16}$$

Again for this special case, the probability of drawing a cell from a given age cohort that exhibits a replicated plasmid follows from evaluation of $1 - p_1(a)$ using Eqs. (2) and (16):

$$p_2(a) = 2^{a/T} - 1. \tag{17}$$

2.2. A titration-handcuffing model for unit-copy-number plasmid replication

Underlying the phenomenological Bernoulli form of the replication rate function (Eq. (12)) is molecular mechanism. In this section, we develop a mathematical model of unit-copy plasmid replication based upon published observations and concepts regarding initiator protein, iteron titration, and handcuffing. Ultimately, it is used to derive a mechanistically based $\lambda(a)$ that is then examined for its ability to reproduce the shape of the Bernoulli form.

2.2.1. General considerations

The model is indicated schematically in Fig. 1. It is based on the production of initiator protein RepE from the *repE* gene under the control of a promotor containing an inverted repeat to which the dimeric form of the initiator (RepE_2) binds to repress its own transcription (Giraldo et al., 2003; Uga et al., 1999). While newly synthesized RepE may be monomer, it strongly dimerizes since the monomer pool is found to be of negligible concentration (Uga et al., 1999). Chaperones convert a portion of the RepE_2 pool to properly folded monomer, RepE_f, which then binds to any of the four 19 bp iteron

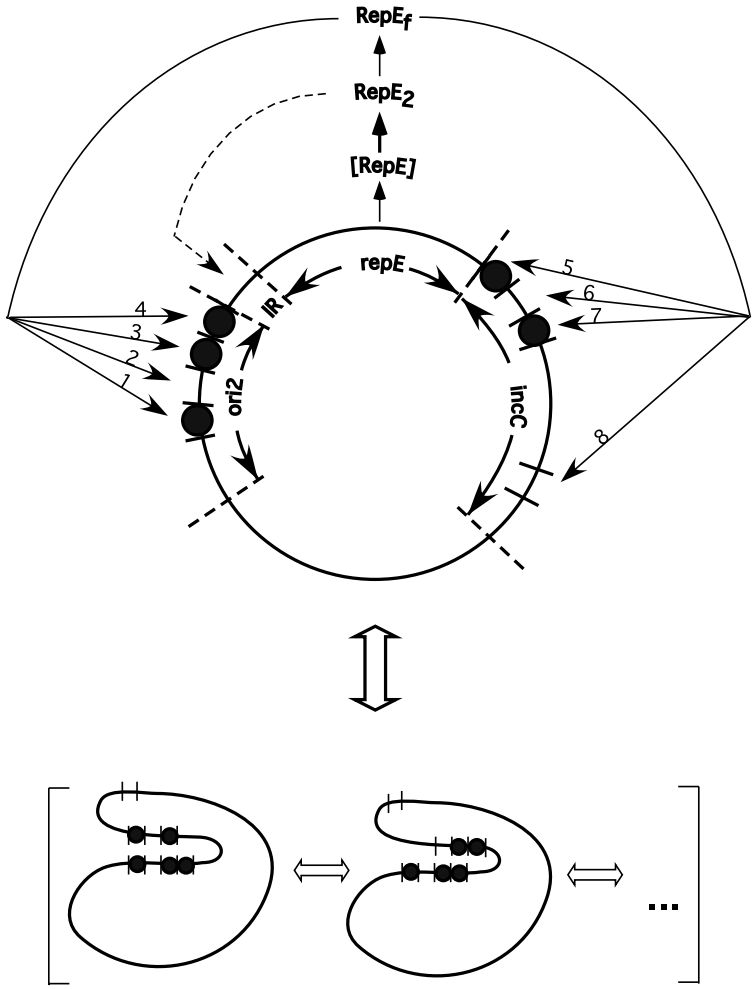

Fig. 1. Molecular components of F-plasmid replication. (A) *repE* gene producing RepE nascent monomer that dimerizes to $RepE_2$, thence progresses via chaperone action to folded $RepE_f$ form. $RepE_f$ then binds randomly to *ori2* iterons (numbers 1–4) and *incC* iterons (numbers 5–8). $RepE_2$ binding to inverted repeat (IR) region (dashed line) inhibits RepE production. Filled circles denote bound $RepE_f$ molecules ($J = 5$). (B) Two possible *cis*-handcuffed forms are shown in equilibrium with the non-handcuffed form above.

sequences in the *ori2* region or to any of the first four iteron sequences in the *incC* region of the plasmid (Giraldo et al., 2003; Komori et al., 1999; Tokino et al., 1986; Uga et al., 1999). A ninth iteron (Iteron 9) in the *incC* region does not bind $RepE_f$ (Uga et al., 1999). When all four of the *ori2* iterons are bound, apparently the origin is remodeled in conjunction with HU and DnaA proteins, resulting in a local melting of the duplex DNA that allows initiation of replication (Kawasaki et al., 1996). Prior to replication, those plasmids with $RepE_f$ bound to iterons in both the *ori2* and *incC* locations may also undergo *cis*-handcuffing, the situation when plasmid DNA folds back on itself such that antiparallel iteron pairing occurs due to contact between bound $RepE_f$ molecules bound to *ori2* and *incC* (Abeles et al., 1995; Chattoraj et al., 1988; Nordstrom, 1990). Handcuffing diverts plasmid configurations from those that might otherwise possess saturated *ori2* iterons and thus reduces the likelihood of replication relative to the case when handcuffing is absent. Finally, once the plasmid has replicated, dilution of the bound initiators over the expanded number of iteron binding sites (titration and *cis*- and

trans-handcuffing are thought to restrict additional rounds of plasmid replication.

2.2.2. Kinetics of RepE initiator protein synthesis

The mathematical representation of this model proceeds from the recognition that the concentration of the dimer $RepE_2$ pool is nearly constant over time because of autorepression and negligible diversion to $RepE_f$ production. This pool consists of 60–100 RepE polypeptides per cell (Uga et al., 1999), and is very similar in magnitude to the value of 40 RepA reported for the P1 plasmid (Swack et al., 1987). During log phase growth, only a minor additional synthesis of $RepE_2$ above that required to make up for volume dilution is needed to replenish the pool converted to monomers by chaperones since a minimum of only about four molecules of $RepE_f$ are required to fully saturate the iteron binding sites over each cell cycle, assuming tight binding. As a consequence, simple simulation (Appendix A) shows that the $RepE_2$ pool concentration fluctuates only by about $\pm 3\%$ over an entire cell cycle.

Accordingly, our basic model treats the $RepE_2$ pool as constant and accounts for the rate of production of $RepE_f$ molecules in a cell by

$$\frac{\mathrm{d}V(a)[\mathrm{RepE}_f(a)]}{\mathrm{d}a} = V(a)k_{\mathrm{chap}}[\mathrm{RepE}_2]. \tag{18}$$

Here $[RepE_f]$ and $[RepE_2]$ are the concentrations of the $RepE_f$ and $RepE_2$ entities, $V(a)$ is the cellular volume as a function of age, and k_{chap} is the rate constant associated with the action of the chaperones. The total number of $RepE_f$ molecules in a single cell is the product of its concentration and the cell volume, $V(a)$ $[RepE_f]$, and the left hand side of Eq. (18) gives its rate of change with age. Defining $V(a)$ $[RepE_f]$ as $J(a)$ and noting that the volume of a cell increases exponentially with age according to $V(a) = V_0 2^{a/T}$ (V_0 is the volume of the cell at birth), Eq. (18) can be rewritten as

$$\frac{\mathrm{d}J(a)}{\mathrm{d}a} = V_0 2^{a/T} k_{\mathrm{chap}}[\mathrm{RepE}_2] = 2^{a/T}k, \tag{19}$$

where all constant factors (including $[RepE_2]$) have been aggregated into a single parameter k. Integration of Eq. (19) yields

$$J(a) = J_0 + \frac{k}{\theta}(\exp[\theta a] - 1), \quad \theta = \frac{\ln 2}{T}, \tag{20}$$

where $J(0)$ has been denoted as J_0. The total number of $RepE_f$ molecules in a cell must exactly double over a cell cycle T if steady state periodicity in the cell population is to be maintained. Hence, the boundary condition $J(T) = 2J_0$ applies. Furthermore, if we denote the gain in $RepE_f$ molecules over each cell cycle as $\Delta J = J(T) - J_0$, then it is apparent from the boundary condition that this gain must equal the initial number of molecules in the cell at birth, i.e.

$$\Delta J = J_0. \tag{21}$$

Replication of a plasmid can occur with certainty only when sufficient $RepE_f$ molecules are generated to guarantee saturation of the four *ori2* iterons. In the absence of handcuffing, this would occur at a minimum value of $J(T) = J_0 + \Delta J = 8$, i.e., when all iterons must be saturated by the end of the cycle. Hence J_0 must equal or exceed 4. In the presence of *cis*-handcuffing, the same minimum value holds if it is assumed that all configurations of binding in which the *ori2* region is saturated represent molecular forms sufficiently remodeled to prevent their participation in *cis*-handcuffing. It is possible for J_0 to exceed the value of 4, but this implies that the plasmids would reach *ori2* saturation at an earlier age and would not be characterized by a replication probability rate exhibiting the rapid monotonic increase near the end of the cell cycle apparent in the functionality of experimentally derived Eqs. (12) or (16). Accordingly, on the basis of this observation, our model employs $J_0 = 4$. Finally, substitution of this value into Eq. (20), together with the observation that $J(T) = 2J_0$, allows the production constant k to be numerically evaluated as $0.0504\,\mathrm{min}^{-1}$.

2.2.3. Plasmid replication rate

As the $RepE_f$ proteins are produced by the chaperones, they bind to the iterons of both the *ori2* and *incC* sets and generate a set of configurations in which the *ori2* iterons are saturated. The probability that an un-replicated plasmid will then undergo replication in the next age increment, $\lambda(a)$, is assumed proportional to the instantaneous

fraction of time spent in this saturated state, numerically equivalent to the fraction of *ori2*-saturated configurations. Thus $\lambda(a)$ is modeled by

$$\lambda(a) = k_\lambda f(J), \tag{22}$$

where $f(J)$ is the instantaneous fraction of *ori2*-saturated configurations corresponding to the presence of J RepE_f molecules in a cell. The proportionality constant k_λ may be interpreted as a constant probability per unit time that an *ori2*-saturated plasmid will undergo the additional interactions and remodeling required for complete initiation of replication.

In the limiting case, where *handcuffing is absent*, and under the assumption of equivalent binding affinity of each iteron, $f(J)$ may be computed from simple combinatorics as the probability of randomly distributing four RepE_f molecules to the four *ori2* sites and the remaining $J-4$ RepE_f molecules to the four binding-capable *incC* sites. From Eq. (B.3) of Appendix B it is

$$f(J) = \frac{\Gamma(J+1)N_{inc}!}{(N_{ori}+N_{inc})!\Gamma(J-N_{ori}+1)}, \quad \text{(no handcuffing)}, \tag{23}$$

where J is the number of RepE_f molecules per cell, $N_{ori}(=4)$ is the number of *ori2* iterons, $N_{inc}(=4)$ is the number of binding-capable *incC* iterons, and the Gamma function, $\Gamma(J+1)$, has been used for $J!$ (Matthews and Walker, 1965).

In the case where *handcuffing is present*, again under the assumption of equivalent RepE_f binding affinity of each iteron, the situation is more complicated and requires an estimation of the $f(J)$ fraction from statistical mechanics (Appendix C). In brief, the partition function q_J is formulated describing all accessible molecular energy states of a plasmid with J initiator molecules bound. Both non-handcuffed and handcuffed states are included in this partition function, including the statistical factors that account for the multiple ways in which molecules may distribute over the iterons leading to molecular configurations of equivalent energy. Those states in which all four *ori2* iterons are saturated are assumed to be sufficiently remodeled into an HU-stabilized loop structure to disallow their participation in handcuffing (Komori et al., 1999). A key approximation that is made in this formulation is to assume that all energy values E_{iJ}^{NHC} which are accessible to the non-handcuffed molecules (NHC) map into a new set of values $\{E_{iJn}^{\mathrm{HC}} = E_{iJ}^{\mathrm{NHC}} + n\Delta E + \varepsilon\}$ when *cis*-handcuffing is present according to the number n of antiparallel RepE_f pair bondings (of incremental energy ΔE) that occur in a particular molecular configuration. The fraction of *ori2*-saturated configurations is then formulated as the ratio of those terms in the partition function corresponding to fully saturated *ori2* configurations divided by the total partition function q_J. The result (Appendix Eq. (C.10)) is

$$f(J) = \frac{m_{0J4}}{m_{0J} + \delta\sum_{n=1}^{3} m_{nJ}\exp(-\beta n\Delta E)}, \quad \text{(handcuffing)}, \tag{24}$$

where m_{0J4} is the number of (non-handcuffed) configurations of J initiator molecules on the iterons consistent with all four of the *ori2* set being bound, m_{0J} is the total number of configurations of J initiator molecules on the iterons in the absence of handcuffing, m_{nJ} $(n=1,3)$ is the number of *cis*-handcuffed configurations (given J) associated with n pair bondings, and β is the inverse Boltzmann constant $\times$ temperature product. δ is a parameter closely related to the reduced degree of spatial freedom available to the handcuffed states relative to the non-handcuffed states. In the limit of infinitely large positive ΔE, and thus, highly unfavored RepE_f pairing, $f(J)$ of Eq. (24) reduces to just m_{0J4}/m_{0J} and is equivalent to the expression derived for the non-handcuffing case, Eq. (23).

The replication rate function $\lambda(a)$ applicable to the handcuffing case, is thus provided by Eq. (22) with $f(J)$ given by Eq. (24). This is a two parameter expression of δ and ΔE provided k_λ of Eq. (22) can be calculated from Eqs. (2) and (13) given an experimental estimate of the incomplete replication fraction ω.

2.2.4. Statistics of pair bonds in cis-handcuffing

The m_{nJ} constants of Eq. (24) depend on structural details of *cis*-handcuffing and may be obtained by the direct counting of possible n pair bond configurations. This count is obtained relatively easily from a first order treatment of

cis-handcuffing in which the four *ori2* iterons of the F plasmid (iterons 1–4) are approximated as equally spaced and opposed in antiparallel orientation by the four binding-capable *incC* iterons (iterons 5–8) in various register. The first three *incC* iterons are also approximated as equally spaced while the fourth one, being substantially distant from the first three, was treated as having a different spacing (Tokino et al., 1986).

Fig. 2 shows a specific example of how the two sets of iterons may register, the number of $RepE_f$ pair bonds generated with each registration, and

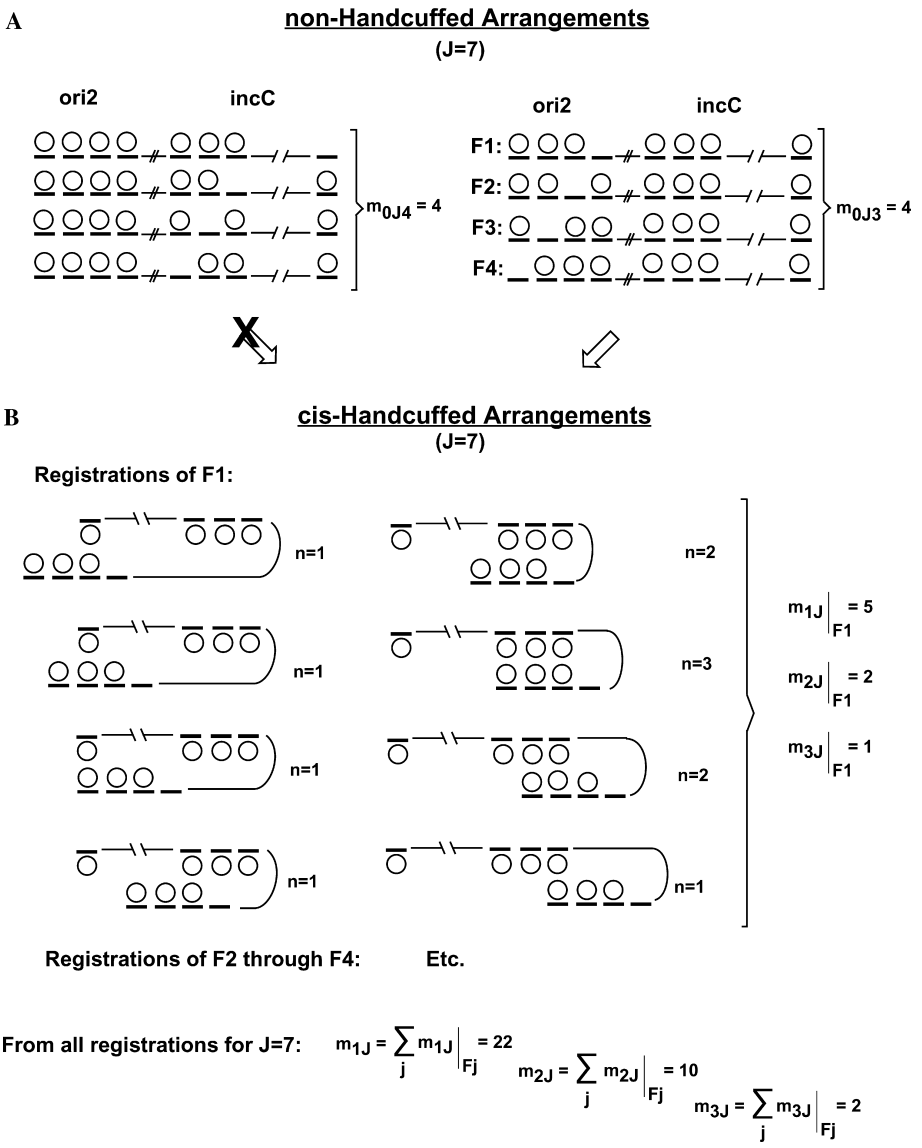

Fig. 2. Statistics of $RepE_f$ pair bonds in *cis*-handcuffing when a total of seven $RepE_f$ are bound ($J = 7$). Non-handcuffed arrangements are shown in (A). Left hand four are *ori2*-saturated forms that are not capable of handcuffing (crossed-out arrow). Right hand four (F1–F4) are capable of handcuffing (intact arrow). (B) Shows all possible *cis*-handcuffed configurations arising from a sliding registration of the form F1. *n* denotes number of pair bonds associated with each registration. The components of m_{nJ} corresponding to the F1 registration are shown at the right hand side of these figures and the m_{nJ} obtained from consideration of all registrations are shown in (B).

configuration counts leading to m_{nJ} when seven initiator molecules in total ($J = 7$) are bound to the plasmid. The upper half of this figure shows all eight possible distributions of these protein molecules without handcuffing. (From Appendix B, this number is denoted by m_{0J} and may be calculated from Eq. (B.2) with $J = 7$.) Note that half of these corresponds to *ori2*-saturated configurations, which because of their remodeling (not shown) are not able to enter into handcuffing. This number of saturated configurations is denoted by m_{0J4} and here equals 4. The remainder of these non-handcuffed forms (denoted by F1 through F4 at top right) is each capable of undergoing handcuffing. Shown in the lower half of Fig. 2 are all of the handcuffed configurations that can be generated from F1 by a sliding linear registration of the *incC* iterons over the *ori2* iterons. The result is five configurations characterized by a single RepE_f pair bond ($n = 1$), 2 by two pair bonds ($n = 2$), and 1 by a triple pair bond ($n = 3$), and these each correspond, respectively, to the F1 contributions to m_{1J}, m_{2J}, and m_{3J} with $J = 7$. Adding similar contributions from the F2, F3, and F4 forms, the final values for m_{1J}, m_{2J}, and m_{3J} with $J = 7$ are found to be 22, 10, and 2. These appear at the bottom of Fig. 2 and, together with the $m_{0J4} = 4$ and $m_{0J} = 8$ values from above, also in the $J = 7$ column of Table 1. The determination of all other m_{0J4} and m_{nJ} values in Table 1 follows from repeated application of the procedure applied in Fig. 2 to the other J values.

Table 1
Number of configurations having n pair bonds given J bound RepE_f molecules[a]

	J				
	4	5	6	7	8
m_{0J4}[b]	1	4	6	4	1
m_{0J}	70	56	28	8	1
m_{1J}	192	240	118	22	0
m_{2J}	8	32	34	10	0
m_{3J}	0	0	2	2	0

[a] Based approximately on iteron spacing of Tokino et al. (1986).

[b] Number of m_{0J} configurations that have a fully bound *ori2* iteron set.

3. Results

3.1. Replication rate function from baby-machine data

The replication rate function, $\lambda(a)$, is given by Eq. (12) (the Bernoulli function) once the slope, A, and inefficiency of replication, ω, have been selected from experimental data. We have selected data corresponding to pML31 miniF plasmids replicating in *Escherichia coli* B/r host cells with a cell division time of 55 min (Fig. 2, panel E of Helmstetter et al., 1997), conditions expected to approximate a unit-copy state of F plasmid. From a log-linear fit to these data, A was determined to be $0.01260\,\mathrm{min}^{-1}$. ω is less well known but has been estimated from the inefficiency in other unit-copy-number plasmid systems as approximately 0.01 (Austin and Eichorn, 1992). For this choice of ω, λ_0 is $0.01247\,\mathrm{min}^{-1}$. (If ω were zero, then only A need to be known and Eq. (17) for $\lambda(a)$ applies.)

The replication rate function so parameterized is displayed as the solid line in Fig. 3A (both main and inset figures). It is apparent that the replication rate is not constant with age and increases monotonically with a rapid upturn toward the end of the cell cycle. The corresponding probability of observing a cell within an age cohort at age a whose plasmid has completed replication, $p_2(a)$, is shown as the solid line in Fig. 3B. This is computed as $1 - p_1(a)$ with $p_1(a)$ given by Eq. (2) and $\lambda(a)$ by Eq. (12). (An insignificantly different curve is generated by the asymptotic limit of Eq. (17).) This result indicates that plasmid replication occurs throughout the cell cycle in spite of the strong monotonic increase of $\lambda(a)$ near the end of the cycle. For comparison, the behavior of $p_2(a)$ with $\lambda(a)$ maintained at a constant initial value of λ_0 is plotted as the dashed line in Fig. 3B.

3.2. Replication rate function from the titration and handcuffing mechanisms

The replication rate function corresponding to the titration-handcuffing model where origin saturation is limiting for initiation is given by Eq. (22), $\lambda(a) = k_\lambda f(J)$, with $f(J)$ in turn given by Eq. (24) when handcuffing is present and by Eq. (23) if

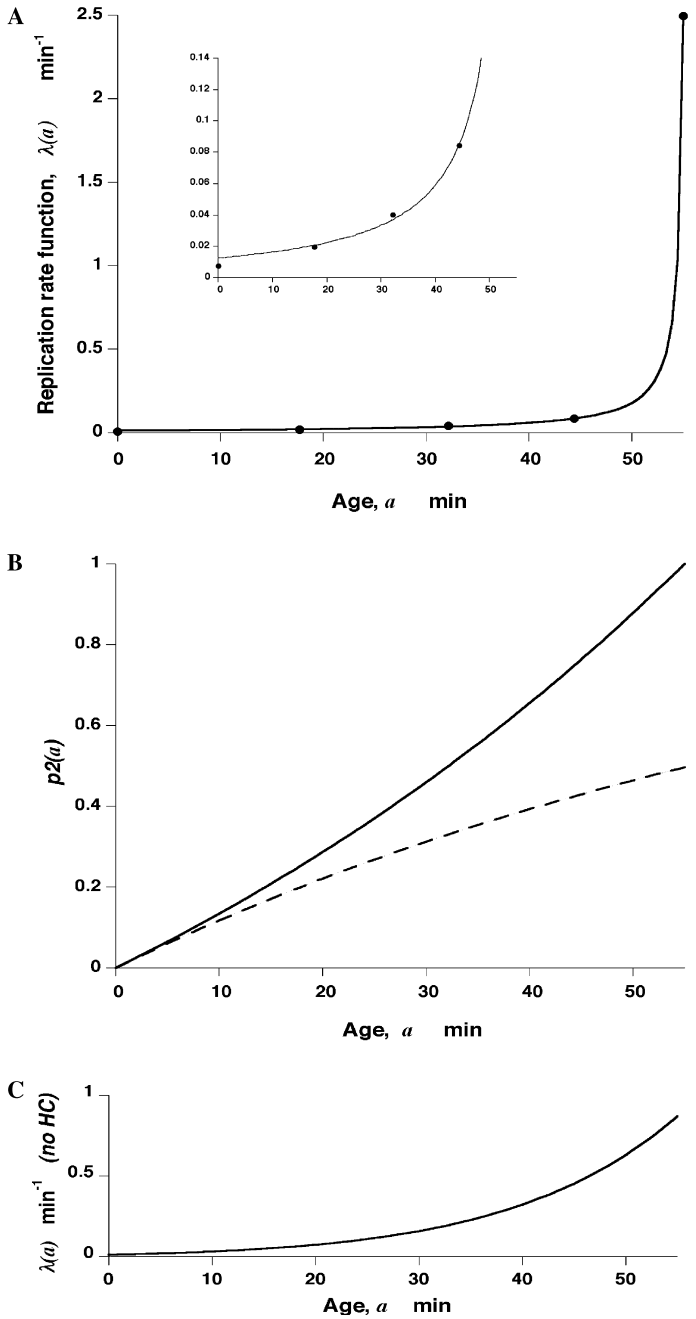

Fig. 3. (A) Replication rate function $\lambda(a)$ as a function of age. Age in minutes, $\lambda(a)$ in min^{-1}. Solid line is Bernoulli function of Eq. (12) derived from baby-machine experiments with $\omega = 0.01$. Points denote theoretical values of $\lambda(a)$ at 0, 18, 32, 44, and 55 min computed from Eqs. (22) and (24) using iteron spacings as published (Tokino et al., 1986). Inset: $\lambda(a)$ at lower ordinate values. (B) Probability $p_2(a)$ of observation of a cell with replicated plasmid. Solid line based on the Bernoulli $\lambda(a)$. Dashed line based on holding the replication rate constant throughout the cell cycle at its initial value of $\lambda_0 = 0.01247\,\text{min}^{-1}$. (C) Replication rate function $\lambda(a)$ when *cis*-handcuffing is absent. From Eqs. (22) and (23) with k_λ chosen so that $\lambda(0) = \lambda_0$. Coordinate scaling is identical to that of (A). Generation time (T) is 55 min in (A–C).

absent. Because mean values of J as a function of age are available from Eq. (20), $f(J)$ in this formula may be re-expressed as a function of age, thus allowing the computation of $\lambda(a)$ as an age-dependent quantity.

A plot of $\lambda(a)$ when handcuffing is absent appears in Fig. 3C. Eq. (23) has been used for $f(J(a))$ and k_λ has been chosen so that $\lambda(0) = \lambda_0$ from the Helmstetter et al. analysis. It is immediately apparent that the curve shape of this replication rate function (particularly at later ages) fails to match the shape of the Bernoulli function obtained from the baby-machine experimentation (cf. Fig. 3A solid line), thus indicating the need to additionally consider the role of handcuffing.

The computation of $\lambda(a)$ for the case when handcuffing is present follows from Eqs. (22) and (24) but requires that estimates be obtained for the unknown parameters δ, β, and ΔE, and k_λ. Of these, δ and $\beta\Delta E$ are the most uncertain ($\beta\Delta E$ is viewed as a single parameter); initially, k_λ can be approximated in terms of these other two parameters by requiring it to satisfy the condition $\lambda(0) = \lambda_0 = k_\lambda f(J(0))$, where λ_0 is the value taken from the baby-machine analysis. To test whether the model, where both titration and handcuffing affect origin saturation, can account for the experimental Bernoulli curve shape, $\lambda(a)$ was evaluated from Eq. (22) at the ages corresponding to integer J via the $f(J(a))$ from Eq. (24) and fit to corresponding values of the Bernoulli expression (Eq. (12) with λ_0 of $0.01247\,\mathrm{min}^{-1}$). From the inverse of Eq. (20) (with $k = 0.0504\ \mathrm{min}^{-1}$), the ages corresponding to integer J (range 4–8) are, respectively, 0, 18, 32, 44, and 55 min. The curve fitting was accomplished by use of the Levenberg–Marquardt implementation of χ^2 minimization, NonlinearRegress, in Mathematica 4.1 (Wolfram, 2001). The m_{nJ} and m_{0J4} values required by Eq. (24) are those of Table 1 derived from the Tokino et al., 1986 spacings. To have available derivatives of the objective function with respect to the parameter variables, each discrete $f(J(a))$ value was replaced by a highly peaked Gaussian centered on the value but spread with a half width of 0.3 min.

Initial fits were conducted with just δ and $\beta\Delta E$ varied. Final fitting allowed k_λ to vary as well, but little variance is associated with this parameter. Final parameter values were found to be: $\delta = 0.34 \pm 0.13$, $\beta\Delta E = -1.32 \pm 0.20$, and $k_\lambda = 2.495 \pm 0.004\,\mathrm{min}^{-1}$. δ is the most uncertain parameter, and only somewhat poorer fits are obtained even for values as low as 0.1 if coupled to a lowered $\beta\Delta E$ nearing -2.0. Also $\beta\Delta E$ is negative as required for energy stabilization of the handcuffed species. A plot of the fit of the values of $\lambda(a)$ from the combination titration and handcuffing model to the experimental Bernoulli curve is given in Fig. 3A. It is apparent that close agreement is achieved between the two evaluations of $\lambda(a)$. The monotonic increase of the replication rate function has been reproduced by the model throughout the age range, including the steep increase toward the end of the cell cycle. The inset shows that agreement is maintained at the lower numerical values of $\lambda(a)$, with only $\lambda(0)$ differing to any significant extent from the Bernoulli value.

To assess the sensitivity of the fit to the assumptions made regarding iteron spacing, another calculation was performed in which all four of the iterons in both the *ori2* and *incC* sets were assumed to be equally spaced from each other. This assumption allows iteron 8 in the *incC* set to more easily enter handcuffing with the iterons of the *ori2* set. After calculating the appropriate m_{nJ} values for this arrangement (Appendix D), a regression to the Bernoulli function revealed an almost identical fit to that displayed in Fig. 3A provided the parameter values shifted to: $\delta = 0.48 \pm 0.16$, $\beta\Delta E = -1.01 \pm 0.20$, and $k_\lambda = 2.495 \pm 0.005\,\mathrm{min}^{-1}$. Hence, the model did not lose its ability to fit the baby-machine derived Bernoulli function with altered spacing assumptions.

4. Discussion

Two principal findings arise from this investigation. The first is that the unit-copy-number plasmid labeling kinetics measured in baby-machine experiments (Helmstetter et al., 1997; Leonard and Helmstetter, 1988) are consistent with a (Bernoulli) replication rate function $\lambda(a)$ that increases monotonically with age and rises

rapidly just before cell division. It is highly suggestive of a replication control mechanism that is designed to force most plasmids to replicate before cells undergo division.

This replication rate function depends on only a few phenomenological assumptions and not on detailed mechanism. Specifically, labeled DNA per cell collected must be log-linear with age, and the plasmid must be unit-copy. The log–linear relationship is well investigated and even seems to hold for the same plasmid when its copy number per cell is changed by varying the growth medium (Helmstetter et al., 1997). In particular, F plasmid replication has conformed to the log-linear fit at generation times between 27 and 90 min. The copy number of plasmids such as P1 and F per cell should vary by a factor of 4–6 under these growth conditions (Austin and Eichorn, 1992). The F plasmid used in baby-machine experiments was of copy-number of about 1.7 and, therefore may not be considered strictly unit copy (Helmstetter et al., 1997). We have assumed that the log–linear replication behavior will not change when the copy number is reduced from 1.7 to one.

The stochastics of replication is described here by a single function $\lambda(a)$ accounting for the probability per unit time of a unit-copy plasmid undergoing replication *conditional* upon having first selected a cell whose plasmid has not yet replicated. As noted above, the rapidly rising replication function does not mean that all replication is compressed into a short period near the end of the cycle (Fig. 3A). This function has significant magnitude even at early ages and consequently allows replication to occur throughout the cell cycle consistent with the observation that replication of the miniF plasmid, pML31, does not occur at a specific point in the cell cycle (Helmstetter et al., 1997). The Bernoulli curve shape of $\lambda(a)$, however, provides an improved description of the probability of replication over the nearly exponentially increasing probability suspected by these authors.

The second principal finding is the ability of a model coupling the concepts of titration and handcuffing to quantitatively account for the replication rate function derived from the baby-machine experiments. Control of replication by initiator titration of iterons alone is insufficient to account for experimental observation (Durland and Helinski, 1990; Pal and Chattoraj, 1988; Tsutsui et al., 1983; Uga et al., 1999). Likewise our model, with handcuffing absent and only titration accounting for the age dependence of $\lambda(a)$ (Eqs. (22) and (23)), also fails to generate the correct curve shape (Fig. 3C). On the other hand, introduction of *cis*-handcuffing into the model for the replication rate does allow agreement with the baby-machine derived function (Fig. 3A). (Since $\lambda(a)$ is a rate based on the conditional probability of observing a cell with unreplicated plasmid, *trans*-handcuffing is not required explicitly in its formulation. *Trans*-handcuffing, however, is implicit as one of the factors restricting multiple rounds of replication.)

The role of *cis*-handcuffing is to shunt RepE_f-bound plasmids (at any particular level of overall binding) to handcuffed forms, so leaving behind a smaller number of the unhandcuffed *ori2*-saturated molecules ready to initiate replication (or, equivalently, reducing the percentage of time a single plasmid spends in the *ori2*-saturated state). As titration of the iterons proceeds (i.e., as J increases), $f(J)$ and hence $\lambda(a)$ increases, but this increase is less rapid when handcuffing is present. (Contrast Fig. 3A with C.) This delay occurs because, progressively, a higher percentage of the available configurations become associated with multiple pair bond states, and the partitioning of molecules toward these states is a highly favorable non-linear function of the number bonds formed.

As presently formulated, our model combining titration and handcuffing evaluates the replication rate function $\lambda(a)$ at discrete ages corresponding to the attainment of integer mean values of bound initiator protein in an aging cohort of cells. While these values lie very close to the baby-machine-derived expression for $\lambda(a)$ and capture the rapid monotonic increase near cell division, linear interpolation between the last two ages (i.e., between $J = 7$ and 8) suggests that the model may be predicting the replication rate function to rise prematurely by as much as 7 min. (Linear interpolation at the age cohort corresponding to $J = 7.5$ implies that the cells in this cohort are

Table 2
List of symbols

a	Age of cell
A	Slope of log-linear $\ln D$ vs a curve
β	Inverse (Boltzmann constant by temperature product)
D	Average plasmid-radioactivity per cell collected
δ	Fitting parameter related to molecular degrees of freedom
Δ	Time interval of thymidine labeling
Δ_e	Time interval of cell collection
ΔE	Incremental energy associated with single RepE pair bond
E_{iJ}	ith energy level associated with a non-handcuffed plasmid bearing J RepE_f molecules
E_{iJn}	ith energy level associated with a plasmid bearing J RepE_f molecules and possessing n pair bonds
ε	Incremental energy change of handcuffing independent of n
$f(J) = f(a)$	Fraction of *ori2*-saturated configurations of a plasmid bearing J RepE_f molecules
G, G_t	Un-repressed and total repE copy number per cell
g_{iJ}	Multiplicity factor for energy state E_{iJ}
g_{iJn}	Multiplicity factor for energy state E_{iJn}
γ	Fixed amount of radiolabel incorporated into a single plasmid's DNA during replication
Γ	Gamma function
h_{iJ}	g_{iJ} exclusive of number of arrangements of J bound proteins on the iterons ($h_{iJ} = g_{iJ}/m_{0J}$)
h_{iJn}	g_{iJn} exclusive of number of arrangements of J bound proteins with n pair bonds on the iterons ($h_{iJn} = g_{iJn}/m_{nJ}$)
HC	Handcuffing
incC	Incompatibility C region of F plasmid
$J(a)$	Total number of RepE_f molecules bound to iterons at age a
J_0	Total number of RepE_f molecules bound to iterons at birth ($a = 0$)
k	Combined rate constant, $V_0 k_{\mathrm{chap}}$ [RepE_2]
k_{chap}	Rate constant of monomer RepE production from dimer by chaperones
k_λ	Probability per unit time of an *ori2*-saturated plasmid initiating replication
$\lambda(a)$	Replication rate function (i.e., probability per unit time of a unit-copy plasmid initiating replication)
λ_0	Value of $\lambda(a)$ at birth
m_{0J}	Total number of configurations of J initiator molecules on iterons in absence of handcuffing
m_{nJ}	Number of *cis*-handcuffed configurations (given J) associated with n pair bondings ($n = 1, 3$)
m_{0J4}	Number of (non-handcuffed) configurations of J initiator molecules on the iterons with the *ori2* set saturated
n	Number of RepE protein pair bonds in a handcuffing configuration
$n(t, a)$	Number density of all cells per age interval at a and t (i.e., total cell density)
$n_1(t, a)$	Number density of *unlabeled* unit-copy cells per age interval at a and t
$n_2^*(t, a)$	Number density of *labeled* 2-copy cells per age interval at a and t
N_{inc}	Number of *incC* iterons (capable of binding RepE_f)
N_{ori}	Number of *ori2* iterons
NHC	No handcuffing
ori2	Replication origin region of F plasmid
ω	Probability of detecting a cell still containing only an unreplicated plasmid at cell division
$p_1(a)$	Probability of selecting a cell within an age cohort centered on age a that contains a single unreplicated plasmid
$p_2(a)$	Probability of selecting a cell within an age cohort centered on age a that contains a replicated plasmid
q_J	Partition function for a plasmid bearing J RepE_f molecules
repE	Gene for RepE initiator protein
RepE	RepE initiator protein (nascent monomer)
RepE_2	Dimer form of RepE
RepE_f	Folded monomer form of RepE capable of iteron binding
σ	Synthesis rate of RepE_2 per un-repressed repE gene
t	Chronological time
t_{ex}	Time of experimental sampling in baby-machine experiments
T	Cell cycle time of host bacterium
θ	$(\ln 2)/T$
$V(a)$	Bacterial cell volume at age a
V_o	Bacterial cell volume at birth ($a = 0$)

evenly divided between $J = 7$ and $J = 8$ cells.) A portion of this early rise (up to 3 min) is attributable to the neglect of the finite times required for thymidine labeling and elution in our derivation of $\lambda(a)$ from baby-machine data. It is possible that the remaining time may be accounted either to a short delay in initiation and labeling after the *ori2*

site becomes saturated, or to a small error in the assumption of the log-linear form Eq. (9) used to fit the experimental data.

A continuous form of $\lambda(a)$ can be derived by expanding the model to include the time-dependent stochastics of $RepE_f$ production and binomial distribution at cell division (leading at any age not just to the mean J as used in the present model but the distribution about this mean) (Morrison et al., 1983). This, however, has been left for future study since the present formulation is sufficient to demonstrate the fundamental finding that a combined titration and handcuffing model appears capable of accounting for the principal shape of the baby-machine-derived $\lambda(a)$.

The combined model that best accounts for the pML31-derived Bernoulli $\lambda(a)$ exhibits a titration of half of the iterons over the cell cycle period, the number of bound $RepE_f$ molecules rising from $J = 4$ to 8 on unreplicated plasmids. We have assumed that $RepE_f$ distributes equally to daughter cells at birth. In reality, some fluctuations ought to occur and plasmids most likely are born with less or more than four bound $RepE_f$. In that case, additional mechanisms may need to be invoked to modulate $RepE_f$ production to adjust for the deviations of J from 4 at birth. For example, two mechanisms are apparently responsible for $RepE_f$ production. One is mediated by chaperones and the other by the iteron DNA itself (Diaz-Lopez et al., 2003). When J is less than 4, higher unbound iteron concentration can help to accelerate $RepE_f$ production, and the reverse would be true when J exceeds 4 at birth. Such feedback mechanisms would ensure homeostasis in $RepE_f$ production in a cell cycle. The titration *cis*-handcuffing model is also not sufficient to explain the failure of copy number increase upon $RepE_f$ oversupply (Uga et al., 1999). A possibility could be that chaperone or some other host factors become limiting when initiators are oversupplied (Ingmer et al., 2001; Uga et al., 1999).

F plasmid replication has also been reported to be cell cycle-specific, although the specificity is less pronounced than *oriC* plasmids (Keasling et al., 1991, 1992). These results are at odds with the experimental observations of Helmstetter and associates analyzed here. The reason for this discrepancy remains unresolved. We chose the Helmstetter data because they were initially more tractable to analysis. To extend the present model to the case of cell cycle-dependent replication additional as yet unspecified control elements must be added to our model.

While issues remain concerning the effects of the stochastics of initiator production and distribution at cell division, as well as details of chaperone behavior, the present characterization of unit-copy plasmid replication appears to account for the principal observations of replication dynamics. This model should thus serve as an initial framework for further characterization of the quantitative aspects of this process (see Table 2).

Appendix A. Small variation of the dimer RepE concentration over the cell cycle

The concentration of dimer RepE (i.e., $RepE_2$), generated by F plasmids in log-phase bacteria, is characterized by little variation over the cell cycle in spite of the doubled gene dosage following plasmid replication. This is demonstrated by the following analysis of the kinetics of $RepE_2$.

The total number of $RepE_2$ molecules in a cell at age a is $V(a)$ $[RepE_2(a)]$, where $V(a)$ is the volume of the cell and $[RepE_2(a)]$ is the dimer concentration in molecules per unit volume at age a. If losses from this pool for turnover or monomer production (estimated as only four molecules per cycle versus the 30–50 needed to make up for volume dilution with cell growth) are negligibly small, then the approximate rate of change of this pool is

$$\frac{\mathrm{d}}{\mathrm{d}a}\left(V(a)[\mathrm{RepE}_2(a)]\right) = \sigma G(a), \tag{A.1}$$

where $G(a)$ is the number of un-repressed repE genes per cell and σ is the synthesis rate of $RepE_2$ molecules per un-repressed gene. The repE promoter is known to be repressed by $RepE_2$ (Uga et al., 1999). Modeling this repression by simple Michaelis–Menten inhibition, $G(a)$ can be re-expressed as

$$G(a) = G_t(a)/(1 + K_a[\mathrm{RepE}_2(a)]), \quad (A.2)$$

where $G_t(a)$ is the total number of repE genes per cell at age a and K_a is the affinity constant for binding of RepE_2 to the promoter site. If bacterial cell volume doubles exponentially with age according to $V(a) = 2^{a/T}V_0$, where V_0 is the volume of the cell at birth, Eq. (A.2) may be rewritten as

$$G(a) = \frac{G_t(a)}{\left(1 + \left(\frac{K_a}{V_o}\right)\left(\frac{1}{2^{a/T}}\right)V(a)[\mathrm{RepE}_2(a)]\right)}. \quad (A.3)$$

Substitution of Eq. (A.3) into Eq. (A.1) yields a differential expression for $V(a)[\mathrm{RepE}_2(a)]$, i.e.

$$\frac{\mathrm{d}}{\mathrm{d}a}(V(a)[\mathrm{RepE}_2(a)]) = \frac{\sigma G_t(a)}{\left(1 + \left(\frac{K_a}{V_0}\right)\left(\frac{1}{2^{a/T}}\right)V(a)[\mathrm{RepE}_2(a)]\right)}. \quad (A.4)$$

$G_t(a)$ is unity from birth up to the average age of plasmid replication, a_{rep}, and 2 from there to cell division at T. The expression for $V(a)[\mathrm{RepE}_2(a)]$ (and hence $[\mathrm{RepE}_2(a)]$) is obtained by integrating Eq. (A.4) from $a = 0$ to arbitrary age a yielding

$$V(a)[\mathrm{RepE}_2(a)] = V(0)[\mathrm{RepE}_2(0)] + \sigma\int_0^a \frac{\mathrm{d}a'}{1 + (K_a/V_0)2^{-a'/T}V(a')[\mathrm{RepE}_2(a')]} \quad (0 \leqslant a \leqslant a_{rep}), \quad (A.5a)$$

$$V(a)[\mathrm{RepE}_2(a)] = V(a_{rep})[\mathrm{RepE}_2(a_{rep})] + 2\sigma\int_{a_{rep}}^a \frac{\mathrm{d}a'}{1 + (K_a/V_0)2^{-a'/T}V(a')[\mathrm{RepE}_2(a')]} \quad (a_{rep} < a \leqslant T). \quad (A.5b)$$

This expression may be evaluated numerically provided estimates are developed for K_a/V_0 and σ. T is known. While the age dependence of $V(a)[\mathrm{RepE}_2]$ also depends on a_{rep}, it will be seen that the *range* of $[\mathrm{RepE}_2]$ is almost independent of the choice for a_{rep} and we have initially given it arbitrary values between 0 and T.

K_a/V_0 can be roughly approximated if it is assumed that the repE promoter is half repressed mid-cycle by the normal RepE_2 concentration of 50 molecules per cell (Uga et al., 1999). In this case, the denominator of Eq. (A.4) must be 2, twice the un-repressed value, and K_a/V_0 is 0.03. σ can be obtained from the periodic boundary condition that applies to Eq. (A.4) under logarithmic growth, namely, the requirement that the total number of RepE_2 molecules must exactly double over the cell cycle (as does the cell volume) or, equivalently, that the RepE_2 concentration is continuous across cell division. Stated mathematically

$$2V(0)[\mathrm{RepE}_2(0)] = V(T)[\mathrm{RepE}_2(T)]. \quad (A.6)$$

Given an estimated initial (birth) number of 33.3 RepE_2 molecules per cell ($= V(0)[\mathrm{RepE}_2(0)]$) and a pre-cell division number of 66.7 ($= V(T)[\mathrm{RepE}_2(T)]$) (Uga et al., 1999), and a choice for a_{rep}, the value of σ is then computed as that which allows Eqs. (A.5a), (A.5b) to satisfy Eq. (A.6).

$[\mathrm{RepE}_2]$ was finally computed as a function of age from Eq. (A.5a), (A.5b), initially for a choice of $a_{rep} \approx T/2$ (from Fig. 3B, the approximate experimental mean replication age of the plasmid). Other parameters were: $T = 55$ min, $K_a/V_0 = 0.03$, and $\sigma = 0.80113$. The maximum and minimum values of $[\mathrm{RepE}_2]$ and its range over the cell cycle are (in units of molecules per cell volume): maximum = 33.33, minimum = 31.45, range = 1.88. The variation over the cell cycle is thus on the order of only $\pm 3\%$. The mean $[\mathrm{RepE}_2]$ value, range, and percent variation remained unchanged for other choices of a_{rep}, only the location of the minimum shifting to a_{rep}. Likewise, other selections for K_a/V_0 led only to compensating changes in σ, and no significant change in the $\pm 3\%$ variation.

Appendix B. Derivation of $f(J)$ in the absence of cis-handcuffing

In the absence of handcuffing, the fraction of *ori2*-saturated plasmids, $f(J)$, is just the number of random arrangements of J initiator molecules ($4 \leqslant J \leqslant 8$) over two groups (the *ori2* and *incC* iteron sets) such that one group contains four molecules (corresponding to a saturated *ori2*

iteron set), divided by the total number of possible arrangements of J molecules over the two groups.

The number of saturated *ori2* arrangements (defined as m_{0J4} for the F plasmid) is equal to the number of combinations of the $J - N_{ori}$ initiator molecules (those remaining after saturation of the N_{ori} *ori2* sites) over the N_{inc} *incC* sites, i.e.

$$m_{0J4} = \frac{N_{inc}!}{(J - N_{ori})!(N_{inc} - J + N_{ori})!}. \tag{B.1}$$

The total number of possible arrangements of the J protein molecules over all of the iterons (defined as m_{0J}) is equal to the number of combinations of J molecules over $N_{ori} + N_{inc}$ sites, i.e.

$$m_{0J} = \frac{(N_{ori} + N_{inc})!}{J!(N_{ori} + N_{inc} - J)!}, \tag{B.2}$$

$f(J)$ is the ratio of these two quantities, i.e.

$$f(J) = \frac{m_{0J4}}{m_{0J}} = \frac{J!N_{inc}!}{(J - N_{ori})!(N_{ori} + N_{inc})!}. \tag{B.3}$$

(Note that no statistical mechanical corrections for temperature effects on the distribution are required under the first order assumption that each plasmid-RepE$_f$ binding energy is independent of RepE$_f$ neighbors and hence that all J-bound plasmids are characterized by the same energy states.)

Appendix C. Derivation of $f(J)$ with cis-handcuffing present

The derivation of the fraction of *ori2*-saturated plasmids, $f(J)$, in the presence of *cis*-handcuffing proceeds from the statistical mechanical partition function q_J for a plasmid bearing J initiator molecules. It may be written as

$$q_J = \sum_i g_{iJ} e^{-\beta E_{iJ}}, \tag{C.1}$$

where E_{iJ} is the ith accessible energy state for a particular molecule with J RepE$_f$ molecules bound to its iterons, and g_{iJ} is a multiplicity factor accounting for quantum degeneracies, as well as multiplicities of E_{iJ} due to the large number of spatial configurations available to any particular molecule and the number of equivalent-energy statistical arrangements of the J-bound molecules on the iteron sites. β is the usual inverse Boltzmann temperature factor. Noting that q_J is composed of terms arising from both the *cis*-handcuffed (HC) and non-*cis*-handcuffed (NHC) forms of the molecules, it may be rewritten as

$$\begin{aligned} q_J &= q_J^{\mathrm{NHC}} + q_J^{\mathrm{HC}} \\ &= \sum_i g_{iJ}^{\mathrm{NHC}} e^{-\beta E_{iJ}^{\mathrm{NHC}}} + \sum_i g_{iJ}^{\mathrm{HC}} e^{-\beta E_{iJ}^{\mathrm{HC}}}. \end{aligned} \tag{C.2}$$

The NHC sum (first right-hand sum) may be further expanded if the set of accessible energy states E_{iJ}^{NHC} is considered to be only negligibly affected by the particular arrangement of the J proteins on the iteron sites. In that case, the multiplicity denoted by g_{iJ}^{NHC} may be considered to arise from a sum over all products of the spatial degeneracy (essentially the number of shapes that a linear-circular polymer may assume in space) for a plasmid with a particular arrangement of n' RepE$_f$ molecules bound to its *ori2* region and $J - n'$ molecules bound to its *incC* region (spatial degeneracy factor h_{iJ}^{HNC}) multiplied by a statistical factor accounting for the total number of $n', J - n'$ arrangements ($m_{0Jn'}$). The q_J^{NHC} term of Eq. (C.2) thus, becomes

$$\begin{aligned} q_J^{\mathrm{NHC}} &= \sum_i \left(\sum_{n'} h_{iJ}^{\mathrm{NHC}} m_{0Jn'} \right) e^{-\beta E_{iJ}^{\mathrm{NHC}}} \\ &= \left(\sum_{n'} m_{0Jn'} \right) \sum_i h_{iJ}^{\mathrm{NHC}} e^{-\beta E_{iJ}^{\mathrm{NHC}}} \\ &= m_{0J} \sum_i h_{iJ}^{\mathrm{NHC}} e^{-\beta E_{iJ}^{\mathrm{NHC}}}, \end{aligned} \tag{C.3}$$

where m_{0J} is given by the last equality and is equal to the total number of available binding configurations for J non-handcuffed proteins.

The fraction of *ori2*-saturated plasmids in the *absence* of handcuffing, $f(J)$, may be derived as the portion of the partition function corresponding to the fully bound *ori2* region (for the F plasmid, the term $m_{0J4} \sum_i h_{iJ}^{\mathrm{NHC}} e^{-\beta E_{iJ}^{\mathrm{NHC}}}$) divided by the entire NHC partition function q_J^{NHC}, i.e.

$$f(J) = \frac{m_{0J4}\sum_i h_{iJ}^{\mathrm{NHC}} \mathrm{e}^{-\beta E_{iJ}^{\mathrm{NHC}}}}{m_{0J}\sum_i h_{iJ}^{\mathrm{NHC}} \mathrm{e}^{-\beta E_{iJ}^{\mathrm{NHC}}}} = \frac{m_{0J4}}{m_{0J}} \quad \text{(no handcuffing)}. \tag{C.4}$$

The HC sum (second right-hand sum of Eq. (C.2) may also be expanded to account for the likelihood of various RepE_f pair bondings in *cis*-handcuffing. To do so, a more detailed *cis*-handcuffing model need be specified. In the F plasmid, the four iterons of the *incC* region are able to fold back and assume partial register with the four *ori2* iterons and establish anywhere from one to three pair bonds between the RepE_f molecules previously bound to the DNA. The energies associated with such *cis*-handcuffed plasmids are necessarily lower than those of their non-handcuffed counterparts, primarily by the energy of interaction generated by the pair bonding. As a first approximation to incorporating this in the HC partition function, we assume that each of the energy states accessible to the non-handcuffed plasmids is decreased by an amount equal to the number of pair bonds (n) in the handcuffed counterpart multiplied by the (constant) energy increment of each bonding (ΔE) plus a constant (ε) reflective of energy changes that do not depend on the exact number of pair bonds but only on the presence of a handcuffing locus, i.e., each of the E_{iJ}^{NHC}energies maps into a set of HC energies $\{E_{iJn}^{\mathrm{HC}}\}, n = 1, 3$, where

$$E_{iJn}^{\mathrm{HC}} = E_{iJ}^{\mathrm{NHC}} + n\Delta E + \varepsilon. \tag{C.5}$$

Introducing these energies into HC partition function, and noting that the multiplicity factor g_{iJ}^{HC} must also be mapped into a $\{g_{iJn}^{\mathrm{HC}}\}$ set to account for the expansion of accessible energy states, q_J^{HC} becomes

$$q_J^{\mathrm{HC}} = \sum_i \sum_n g_{iJn}^{\mathrm{HC}} \mathrm{e}^{-\beta E_{iJn}^{\mathrm{HC}}} = \sum_i \sum_n g_{iJn}^{\mathrm{HC}} \mathrm{e}^{-\beta E_{iJ}^{\mathrm{NHC}} - \beta\varepsilon} \mathrm{e}^{-\beta n\Delta E}. \tag{C.6}$$

The g_{iJn}^{HC} multiplicity may then be considered as product of a factor accounting for the spatial degeneracy of a particular *cis*-handcuffed molecule with n bond pairings (h_{iJn}^{HC}) multiplied by the statistical number of n-bond arrangements possible ($m_{nJ}, n = 1, 3$), so that

$$q_J^{\mathrm{HC}} = \sum_i \sum_n h_{iJn}^{\mathrm{HC}} m_{nJ} \mathrm{e}^{-\beta E_{iJ}^{\mathrm{NHC}} - \beta\varepsilon} \mathrm{e}^{-\beta n\Delta E}. \tag{C.7}$$

Further rearrangement is possible if it is observed that h_{iJn}^{HC} depends only weakly on n since the spatial degeneracy is determined more by the large scale figure–eight nature of the handcuffed molecule than by the small scale details of the handcuffing link related to n. Accordingly, the n-subscript may be dropped from h_{iJn}^{HC} and Eq. (C.7) may be regrouped to yield the product

$$q_J^{\mathrm{HC}} = \left(\sum_i h_{iJ}^{\mathrm{HC}} \mathrm{e}^{-\beta E_{iJ}^{\mathrm{NHC}} - \beta\varepsilon}\right)\left(\sum_n m_{nJ} \mathrm{e}^{-\beta n\Delta E}\right). \tag{C.8}$$

This becomes the right hand term in $q_J = q_J^{\mathrm{NHC}} + q_J^{\mathrm{HC}}$.

The fraction of *ori2*-saturated plasmids in the *presence* of handcuffing, $f(J)$, may be derived as the portion of the partition function corresponding to the fully bound, non-handcuffed *ori2* region (from the NHC portion of q_J for the F plasmid, the term $m_{0J4}\sum_i h_{iJ}^{\mathrm{NHC}} \mathrm{e}^{-\beta E_{iJ}^{\mathrm{NHC}}}$) divided by the entire partition function q_J, i.e.

$$f(J) = \frac{m_{0J4}\sum_i h_{iJ}^{\mathrm{NHC}} \mathrm{e}^{-\beta E_{iJ}^{\mathrm{NHC}}}}{q_J^{\mathrm{NHC}} + q_J^{\mathrm{HC}}} = \frac{m_{0J4}\sum_i h_{iJ}^{\mathrm{NHC}} \mathrm{e}^{-\beta E_{iJ}^{\mathrm{NHC}}}}{m_{0J}\sum_i h_{iJ}^{\mathrm{NHC}} \mathrm{e}^{-\beta E_{iJ}^{\mathrm{NHC}}} + \left(\sum_i h_{iJ}^{\mathrm{HC}} \mathrm{e}^{-\beta E_{iJ}^{\mathrm{NHC}} - \beta\varepsilon}\right)\left(\sum_n m_{nJ} \mathrm{e}^{-\beta n\Delta E}\right)}, \tag{C.9}$$

where substitutions for q_J^{NHC} and q_J^{HC} have been taken from Eqs. (C.3) and (C.8). Finally by dividing both numerator and denominator of Eq. (C.9) by the sum appearing in the numerator, and defining the ratio $\sum_i h_{iJ}^{\mathrm{HC}} \mathrm{e}^{-\beta E_{iJ}^{\mathrm{NHC}} - \beta\varepsilon} / \sum_i h_{iJ}^{\mathrm{NHC}} \mathrm{e}^{-\beta E_{iJ}^{\mathrm{NHC}}}$ as a parameter δ, $f(J)$ is obtained as

$$f(J) = \frac{m_{0J4}}{m_{0J} + \delta\sum_{n=1}^{3} m_{nJ} \mathrm{e}^{-\beta n\Delta E}} \quad \text{(handcuffing)}. \tag{C.10}$$

Assuming the m_{0J4} and m_{nJ} factors are available from straightforward combinatorics, Eq. (C.10) is thus a two parameter description of $f(J)$.

Appendix D

Number of configurations having n pair bonds given J bound $RepE_f$ molecules and equal spacing between iterons in both the *ori2* and *incC* sets

	J				
	4	5	6	7	8
$m_{0,J4}$[a]	1	4	6	4	1
$m_{0,J}$	70	56	28	8	1
$m_{1,J}$	221	180	70	10	0
$m_{2,J}$	5	56	52	10	0
$m_{3,J}$	0	0	6	6	0

[a] Number of $m_{0,J}$ configurations that have a fully bound *ori2* iteron set.

References

Abeles, A.L., Reaves, L.D., Youngren-Grimes, B., Austin, S.J., 1995. Control of P1 plasmid replication by iterons. Mol. Microbiol. 18, 903–912.

Austin, S.J., Eichorn, B.G., 1992. Random diffusion can account for *topA*-dependent suppression of partition defects in low-copy-number plasmids. J. Bacteriol. 174, 5190–5195.

Chattoraj, D.K., 2000. Control of plasmid DNA replication by iterons: no longer paradoxical. Mol. Microbiol. 37, 467–476.

Chattoraj, D.K., Mason, R.J., Wickner, S.H., 1988. Mini-P1 plasmid replication: the autoregulation-sequestration paradox. Cell 52, 551–557.

Diaz-Lopez, T., Lages-Gonzalo, M., Serrano-Lopez, A., Alfonso, C., Rivas, G., Diaz-Orejas, R., Giraldo, R., 2003. Structural changes in RepA, a plasmid replication initiator, upon binding to origin DNA. J. Biol. Chem. 278, 18606–18616.

Durland, R.H., Helinski, D.R., 1990. Replication of the broad-host-range plasmid RK2: direct measurement of intracellular concentrations of the essential TrfA replication proteins and their effect on plasmid copy number. J. Bacteriol. 172, 3849–3858.

Giraldo, R., Fernandez-Tornero, C., Evans, P.R., Diaz-Orejas, R., Romero, A., 2003. A conformational switch between transcriptional repression and replication initiation in the RepA dimerization domain. Nat. Struct. Biol. 10, 565–571.

Helmstetter, C.E., Thornton, M., Zhou, P., Bogan, J.A., Leonard, A.C., Grimwade, J.E., 1997. Replication and segregation of a miniF plasmid during the division cycle of *Escherichia coli*. J. Bacteriol. 179, 1393–1399.

Ingmer, H., Miller, C., Cohen, S.N., 2001. The RepA protein of pSC101 controls *Escherichia coli* cell division through the SOS response. Mol. Microbiol. 42, 519–526.

Kawasaki, Y., Matsunaga, F., Kano, Y., Yura, T., Wada, C., 1996. The localized melting of mini-F origin by the combined action of the mini-F initiator protein (RepE) and HU and DnaA of *Escherichia coli*. Mol. Gen. Genet. 253, 42–49.

Keasling, J.D., Palsson, B.O., Cooper, S., 1991. Cell-cycle-specific F plasmid replication: regulation by cell size control of initiation. J. Bacteriol. 173, 2673–2680.

Keasling, J.D., Palsson, B.O., Cooper, S., 1992. Replication of mini-F plasmids during the bacterial division cycle. Res. Microbiol. 143, 541–548.

Kittell, B.L., Helinski, D.R., 1991. Iteron inhibition of plasmid RK2 replication in vitro: evidence for intermolecular coupling of replication origins as a mechanism for RK2 replication control. Proc. Natl. Acad. Sci USA 88, 1389–1393.

Komori, H., Matsunaga, F., Higuchi, Y., Ishiai, M., Wada, C., Miki, K., 1999. Crystal structure of a prokaryotic replication initiator protein bound to DNA at 2.6 angstrom resolution. EMBO J. 18, 4597–4607.

Leonard, A.C., Helmstetter, C.E., 1988. Replication patterns of multiple plasmids coexisting in *Escherichia coli*. J. Bacteriol. 170, 1380–1383.

Matthews, J., Walker, R.L., 1965. Mathematical Methods of Physics. WA Benjamin Inc., New York.

Morrison, P.F., Aroesty, J., Creekmore, S.P., Barker, P.E., Lincoln, T.L., 1983. A preliminary model of double-minute-mediated gene amplification. J. Theor. Biol. 104, 71–91.

Mukherjee, S., Patel, I., Bastia, D., 1985. Conformational changes in a replication origin induced by an initiator protein. Cell 43, 189–197.

Nordstrom, K., 1990. Control of plasmid replication—How do DNA iterons set the replication frequency. Cell 63, 1121–1124.

Pal, S.K., Chattoraj, D.K., 1988. P1 plasmid replication: initiator sequestration is inadequate to explain control by initiator-binding sites. J. Bacteriol. 170, 3554–3560.

Swack, J.A., Pal, S.K., Mason, R.J., Abeles, A.L., Chattoraj, D.K., 1987. P1 plasmid replication: measurement of initiator protein concentration in vivo. J. Bacteriol. 169, 3737–3742.

Tokino, T., Murotsu, T., Matsubara, K., 1986. Purification and properties of the mini-F plasmid-encoded E protein needed for autonomous replication control of the plasmid. Proc. Natl. Acad. Sci. USA 83, 4109–4113.

Tsutsui, H., Matsubara, K., 1981. Replication control and switch-off function as observed with a mini-F factor plasmid. J. Bacteriol. 147, 509–516.

Tsutsui, H., Fujiyama, A., Murotsu, T., Matsubara, K., 1983. Role of nine repeating sequences of the mini-F genome for expression of F-specific incompatibility phenotype and copy number control. J. Bacteriol. 155, 337–344.

Uga, H., Matsunaga, F., Wada, C., 1999. Regulation of DNA replication by iterons: an interaction between the *ori2* and *incC* regions mediated by RepE-bound iterons inhibits DNA replication of mini-F plasmid in *Escherichia coli*. EMBO J. 18, 3856–3867.

Wolfram, S., 2001. Mathematica, Version 4.1. Addison-Wesley Publishing Co., New York.

Communicated by K. Nordström